IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received August 23, 2020, accepted October 8, 2020, date of publication October 19, 2020, date of current version October 29, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3032229

Vibrotactile Captioning of Musical Effects in Audio-Visual Media as an Alternative for Deaf and Hard of Hearing People: An EEG Study

MARÍA J. LUCÍA[1,2], PABLO REVUELTA[2,3], ÁLVARO GARCÍA[4], BELÉN RUIZ[1,2], RICARDO VERGAZ[4], VÍCTOR CERDÁN[5], AND TOMÁS ORTIZ[6]

[1]Department of Computer Science, Carlos III University of Madrid, 28911 Leganés, Spain
[2]Spanish Center for Captioning and Audio Description, Carlos III University of Madrid, 28903 Leganés, Spain
[3]Department of Computer Science, Oviedo University, 33204 Gijón, Spain
[4]GDAF-UC3M, Department of Electronic Technology, Carlos III University of Madrid, 28911 Leganés, Spain
[5]Department of Applied Communication Sciences, Complutense University of Madrid, 28040 Madrid, Spain
[6]Department of Psychiatric, Complutense University of Madrid, 28040 Madrid, Spain

Corresponding author: María J. Lucía (maluciam@inf.uc3m.es)

This work was supported in part by the Comunidad de Madrid through the SINFOTON2-CM Research Program under Grant S2018/NMT-4326-SINFOTON2-CM.

ABSTRACT Standard captioning for the deaf and hard of hearing people cannot transmit the emotional information that music provides in support of the narrative in audio-visual media. We explore an alternative method using vibrotactile stimulation as a possible channel to transmit the emotional information contained in an audio-visual soundtrack and, thus, elicit a greater emotional reaction in hearing-impaired people. To achieve this objective, we applied two one-minute videos that were based on image sequences that were unassociated with dramatic action, maximizing the effect of the music and vibrotactile stimuli. While viewing the video, using EEG we recorded the brain activity of 9 female participants with normal hearing, and 7 female participants with very severe and profound hearing loss. The results show that the same brain areas are activated in participants with normal hearing watching the video with the soundtrack, and in participants with hearing loss watching the same video with a soft and rhythmic vibrotactile stimulation on the palm and fingertips, although in different hemispheres. These brain areas (auditory cortex, superior temporal cortex, medial frontal cortex, inferior frontal gyrus, superior temporal pole and insula) have been consistently reported as areas involved in the emotional perception of music. We conclude that vibrotactile stimuli can generate cortex activation while watching audio-visual media in a similar way to sound. Thus, a further in-depth study of the possibilities of these stimuli can contribute to an alternative subtitling channel for enriching the audiovisual experience of hearing-impaired people.

INDEX TERMS Music emotion recognition, hearing impairment, vibrotactile, audio-visual, captions, accessibility, electroencephalography.

I. INTRODUCTION

Although in audio-visual media images predominate over sounds in case of ambiguity or emotional conflict [1], the capacity of music to generate emotions is widely used in audio-visual media [2] as a support to the narrative [3] and to modulate the emotional tone of a scene [4]. The role of music producing emotional responses is one of its defining features [5], and it is consistent across cultures [6], [7].

However, hearing-impaired people cannot access the information conveyed by music. The rights established in the UN Convention on the Rights of Persons with Disabilities [8] concerning access to television programs, films, theatre, and other cultural activities, focus on captioning. Captioning is the reference assistive tool for hearing impairment. Special regulations have been issued to guarantee its application (in Spain, the General Law of Audio-visual Communication [9]), and its quality, considering factors as visual aspects, synchronism, presentation speed, speaker identification or accuracy [10]. In the specific case of music, it must be subtitled in the upper right of the screen, with a small text in brackets summarizing the type of music, sensation transmitted, or identification of the piece, e.g.: "(Lullaby Plays)." But this textual representation does not transmit the emotional

The associate editor coordinating the review of this manuscript and approving it for publication was Sung Chan Jun.

information contained in the soundtrack; on the contrary, captions produce higher activation of voluntary attentional circuits [11].

Recent studies are increasingly finding analogies between the perceptive mechanisms of touch and hearing. Both senses respond to mechanical energy. There are four types of specialized tactile receptors in the glabrous skin: Ruffini, Merkel, Meissner, and Pacini corpuscles. These last receptors are found in the deepest part of the dermal layer, being very fast adapting and extremely sensitive to tactile vibration, activated with frequencies between 50 and 1000 Hz, with a maximum sensitivity around 200-300 Hz. The Meissner corpuscles are superficial, fast adapting, and their maximum sensitivity to tactile vibrations is in the frequency range between 5 and 50 Hz [12]. These two types of receptors have a low activation threshold and can be activated by mechanical skin deformation of the order of 100 Angstrom [13]. Together, they cover a range of 5-1000 Hz while the total auditory range is 20-20000 Hz. Different studies have focused on the parameters of vibrotactile perception. For example, it has been shown that the vibrotactile pitch range of musical notes from C1 to G5 can be discriminated, but that intervals of less than three semitones are not detected [14]. A success rate above chance in discriminating musical timbres based on vibrotactile stimulation alone has been found [15], and that subjects can distinguish between different pure and complex waveforms [16]. Although the neurological mechanisms that encode information about the tactile vibration are not yet clearly identified, Kayser *et al.* suggest that there is an area of vibrotactile-auditory coactivation [17], and that auditory-tactile integration occurs early and close to primary sensory areas. Evidence of vibrotactile input to the human auditory cortex has also been recorded with low frequency vibrations detected by both tactile and auditory systems [18]. Huang *et al.* link the perception of vibrotactile rhythm to the activation of the Pacini receptors, which encode the temporal patterns of vibrations in a similar way to the encoding of sound waves in the cochlea, and to the same perceptual mechanism common to both senses [19]. Concerning timbre, Russo *et al.* report that, analogous to auditory discrimination of timbre, the vibrotactile discrimination could be the result of the integration at the cortical level of the outputs of tuned tactile receptors that would work as bandpass filters [15]. Interactions between audition and touch in frequency and temporal perception have been found in different studies [19]–[21]. Rahman *et al.* showed, using fMRI, that the temporal frequency information contained in audio-tactile stimuli was processed by feature-specific somatosensory and auditory systems [22].

In the last decade, the study of the relationship between musical features and emotion has gained much interest in the field of neuroscience [23], [24]. Mode and tempo are the parameters that most influence the type of emotion generated by music together with register, dynamics, articulation, and timbre, although it may vary across emotions [23]. The challenge is to find how these musical parameters can be effectively mapped to a vibrotactile pattern to produce the same emotional effects. Different types of devices, as vibrating armchairs, vests, bracelets, feet shakers, have already been designed to map the acoustic musical signal to a vibrotactile one, to improve experiences of music, musical practice, or musical performance, especially for hearing-impaired people [25]–[27]. Some approaches to the treatment of the acoustic signal for its transformation into a tactile signal have been tested, generally considering that it may not be necessary to encode all the auditory information available in the tactile channel. As the tactile frequency perception range is 5-1000 Hz versus 20-20000 Hz for the ear, it is easier to feel the sound with a large bass component, by using low-pass filters, or by removing harmonics from the fundamental frequency [28]. User feedback has been very positive when rhythmic music is translated into tactile vibration, while more harmonic music, with less marked rhythms, have been shown to produce sensations of buzz or diffused vibration [26], [28], [29]. One promising approach to transmit tone information is to encode it in a spatial tactile dimension, by applying different frequency bands in different areas of the skin [28], [29].

In their review article about affective sound processing, Frühholz *et al.* compile the brain regions that are consistently reported to be involved in musical affective processing [30]. Among these zones, the ones that can be registered by EEG are auditory cortex, superior temporal cortex, medial frontal cortex, inferior frontal gyrus, and insula. The auditory cortex seems to integrate emotional information decoding complex sounds features that evolve in time [30]. The superior temporal cortex is responsible for multimodal information integration [31], [32]. The insula would be involved in the translation of the auditory perception of sounds into a self-experience of emotion [33], [34]. The medial frontal cortex supports several social and emotional functions related to interpersonal communication and understanding [35] and possibly links music to memory-associated emotional processes [36], [37]. The inferior frontal gyrus is involved in adaptive response preparation in a perception-action cycle [30].

User feedback on feeling music through vibrotactile devices has been measured, mostly based on self-reports in the context of musical settings. We present here an experimental work based on the next hypotheses:

- a rhythmic vibrotactile stimulation, applied to people with profound and severe loss of hearing while experiencing audio-visual media, can elicit an emotional response in regions of the brain like those produced by the music soundtrack in normal-hearing participants.
- the same brain areas can be activated both by this simple vibrotactile stimuli in hearing-impaired participants and by music in normal-hearing participants, while watching the same videos.

We used EEG records to measure and compare the brain areas activated by the vibrotactile stimulation in hearing-impaired participants, with the brain areas activated by music in normal-hearing participants while watching the same videos.

The vibrotactile stimulation was a series of 102 ms square-wave vibrations of 1 kHz based on a rhythmic pattern, as rhythm is the musical parameter that has been reported to have better user feedback when translated into tactile vibration [26].

II. MATERIALS AND METHODS

A. PARTICIPANTS

Two groups of participants were recruited. For the control group, volunteers were solicited from the medical school where the study was conducted. For the experimental group, volunteers were requested through associations of deaf people. All volunteers were informed of the objective and general procedure of the study.

For the control group, 9 female participants, with self-reported normal-hearing, were recruited, aged between 18 and 22 (mean: 19.22, SD: 1.31), all of them had a high school degree and were medical students at the Complutense University of Madrid.

For the experimental group, 7 female participants with self-reported hearing loss were recruited, aged between 37 and 61 (mean: 49.28, SD: 10.76), one with a high school degree, 4 with a college degree, and 2 with post-graduate studies.

Only female participants were considered to eliminate the genre variable. The genre has been associated with significant brain differences in the way emotional stimuli are processed. For example, unpleasant and highly arousing stimuli evoke greater ERP amplitudes in women than in men [38], or stimuli considered beautiful by the participants produced bilateral brain activity in women, while in men only in the right hemisphere [39].

The self-reported hearing losses were classified according to the Audiometric Classification of Hearing Impairments of the International Bureau for Audiophonology [40].:

- mild hearing loss (between 20 and 40 dB)
- moderate hearing loss (between 41 and 70 dB), where speech is perceived if the voice is loud, and the subject understands better what is being said if he can see his/her interlocutor)
- severe hearing loss (between 71 and 90 dB, where speech is perceived if the voice is loud and close to the ear, and loud noises are also perceived)
- very severe hearing loss (between 91 and 119 dB, where speech is not perceived, and only very loud noises are perceived)
- total hearing loss (over 120 dB).

Four participants had very severe hearing loss and used hearing aids (one of them had a cochlear implant in one ear), and three participants had total hearing loss (one of them had a cochlear implant). All participants were right-handed.

Finally, an informed consent document was signed, as approved by the Bioethical Committee of the Carlos III University of Madrid, and a survey was filled out concerning demographic information, level of studies, and degree of hearing loss.

B. MATERIALS

1) STIMULI

Two videos of one-minute length were created, based on sequences of images that were unassociated with a dramatic action, given that images must be kept as neutral as possible to allow the measurement of the music and vibrotactile effects, according to [41], [42]. One video displayed aerial images of cereal fields at sunset. The other video showed images of organisms and plants moving in the depths of the sea. The sequences -only video tracks- were extracted from *The Straight Story* (David Lynch, 1999) and *The Tree of Life* (Terrence Malick, 2011).

Two orchestral musical fragments were selected based on their assessed relation with the evocation of emotions: an excerpt from Nicolo Paganini, Violin Concerto No. 1, Op. 6, Rondo, with fast tempo (106 bpm), and major scale, and an excerpt from Samuel Barber, Adagio for String, Op.11, with slow tempo (60 bpm), and minor scale. These stimuli have been successfully used to elicit emotions previously [43], [44].

Two vibrotactile stimuli were programmed on a glove, specifically designed for the experiment, which applied a soft tactile vibration on the fingertips and the palm. The glove has been designed choosing stimuli locations with a psycho-physical AFC (alternative forced-choice) method [45]. A glove was selected due to comfort and easy installation of the stimulation motors. Stimuli were produced at the bpm of the musical excerpts, synchronized with them, each beat lasting 102 ms.

The videos were organized in 3 different conditions. In condition 1, the video soundtrack was randomly associated with one of the 2 musical fragments. In condition 2, the videos had no soundtrack. In condition 3, the videos had no soundtrack and were randomly associated with one of the 2 vibrotactile stimuli.

2) HARDWARE

In this experiment, two computers were involved. The first triggered the video, sending temporal marks to the EEG amplifier to locate the stimuli and sending the synchronization signal to the glove, with screen and speakers pointing to the participant. The second computer registered the EEG data. EEG data were recorded with a multichannel EEG equipment (64 channels), using a custom-designed electrode Neuroscan cap and an ATI EEG system (Advantek SRL). Additional channels were included to monitor eye movement (right and left lateral canthi and superior and inferior orbits of the left eye). The reference electrodes were placed on the mastoids, and the ground electrode was placed on the forehead. Data were processed to an average reference following acquisition with a band-pass filter of 0.05–30 Hz at a sample rate of 1000 Hz. Impedances were kept under 5 kΩ.

Segmentation of EEG trials related to each subject was performed by using a window of duration 1 s over each

FIGURE 1. Glove used for vibrotactile stimulation.

1-minute condition. Individual subject recordings were visually inspected to ensure that clean recordings were obtained. Eye and muscle movement artifacts were identified off-line on a trial-by-trial basis through visual inspection. The trials which involved artifacts like eye movements or eye blinks were rejected. From the remaining artifact-free trials, averages were computed separately for each participant and condition. Only the grand averages were considered to estimate the differentially activated regions, regardless of the frequency bands involved.

Low-resolution electromagnetic tomography (LORETA) [46] was applied to the individual averaged recordings to identify the underlying neural current sources of scalp potentials. LORETA is a reverse solution method that calculates the three-dimensional distribution of neural generators in the brain as a current density value (A/m2) for a total of 2.394 voxels. The method used in this work considers different anatomical restrictions taken from a 90-part segmentation of the average brain atlas of the Montreal Neurological Institute (MNI) [47]. Finally, for all subjects in the same condition the average source location map was obtained.

For vibrotactile stimulation, a glove was used (Fig. 1). This subsystem was implemented on an Inesis Golf Glove 100, with both the size and elastic fabric ensuring that the glove fit the participants' fingertips and palm. Coin motors Uxcell 1030 were placed to provide the proper skin stimulation on each of three specific locations: index finger, ring finger, and palm. These motors are powered with 3 volts DC and 70 milliamps and were selected due to their small size. Motors were glued to the outside of the glove, and properly transmitted the vibrations to the skin due to the fit. Before each test, a researcher checked the fit of the glove, and whether the participants were comfortable with it and could clearly feel the vibrations generated.

The motors driving signal was a 1 kHz square signal burst of 102 ms length, generated by an Arduino UNO triggered by the control PC, and synchronized with the music. 1 kHz frequency was used to produce a fast activation of the motor vibration, as it is its resonant frequency. 102 ms is time enough to guarantee that the participant feels a burst. The rhythm at which the activations were triggered was the main stimulus, i.e. the music tempo.

C. PROCEDURE

Participants were invited to attend individual sessions. They were first asked to fill in a survey including questions about their age, gender, education level, type and degree of hearing loss and hearing aids, if applicable, and dominant hand.

The experiment took place in a soundproofed room. Participants were asked to sit on a comfortable sofa and face a screen of 25 inches, embedded in the wall, that was at 1.5 m in direct line of sight.

The sound was provided, to hearing participants, through stereo speakers attached with the screen. Hearing loss participants were asked to remove their hearing aids, but to keep their glasses on if needed. The EEG cap was fixed to their head and the vibrotactile stimulation was provided through the right hand (as there were no left-handed participants recruited).

It was explained to the participants that they were going to watch a video and they were instructed to sit comfortably and relax during the video projection. Lights in the room were turned off, and the corresponding video was played. Audio tracks were not played to hearing-impaired participants to avoid interferences of eventual hearing rests detected in previous experiments.

There were three conditions for the experiment:

- Condition 1: control group participants watched the 2 videos in random order, and each video with one of the two musical fragments randomly selected as its soundtrack.
- Condition 2: experimental group watched the 2 videos in mute mode and random order.
- Condition 3: experimental group watched again one of the 2 videos, randomly selected, in mute mode, while simultaneously the vibrotactile glove produced one of the 2 vibrotactile stimuli tracks (corresponding to one of the two music fragments) randomly selected.

Both positive and negative emotional audio stimuli, and their vibrotactile correlates, were used to calculate the grand average of all reactions, eliminating the valence effect. In the hearing-impaired participants, only one of the emotional vibrotactile stimulation was used to reduce the habituation effects to such stimuli.

The control group, hearing participants, were shown the video with sound, as reference to a complete audiovisual experience with which to compare the experimental group.

III. RESULTS

To compare the three conditions, we considered the ten maximum statistically significant activation peaks in each group obtained with the Hotelling's T^2 test against zero. The Hotelling's T^2 test is a multivariate statistical analysis for hypothesis testing based on covariance matrix. Applied on the EEG registered values over all the electrodes, in neuroscientific experiments, it has proven to detect brain areas with statistically significant activation compared with the average state.

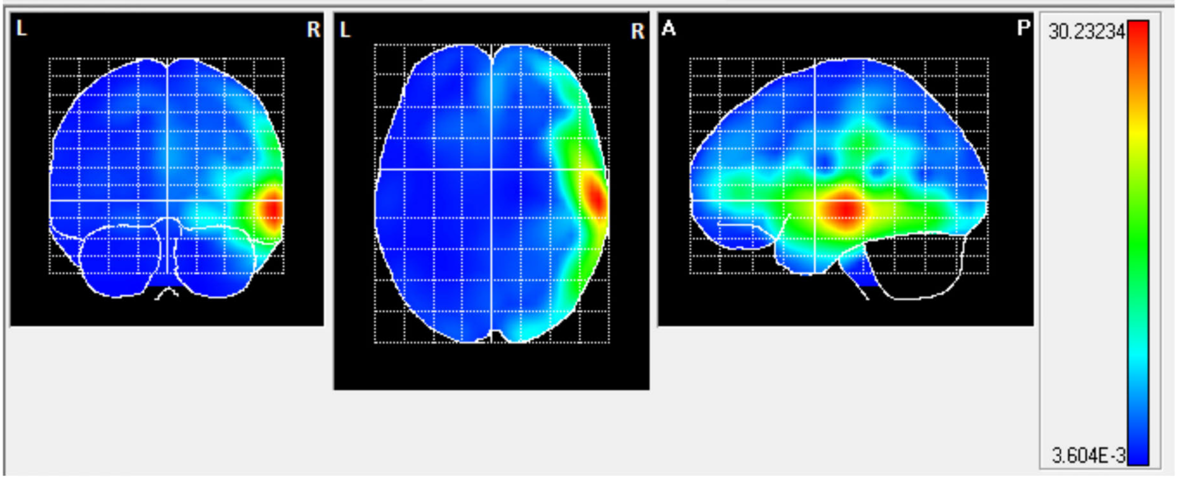

FIGURE 2. | Mean electrical maps for Condition 1 (hearing participants, video + music). Maximal intensity projection areas are displayed in yellow/red color. SPMs were computed based on a voxel-by-voxel Hotelling's T^2 test against zero.

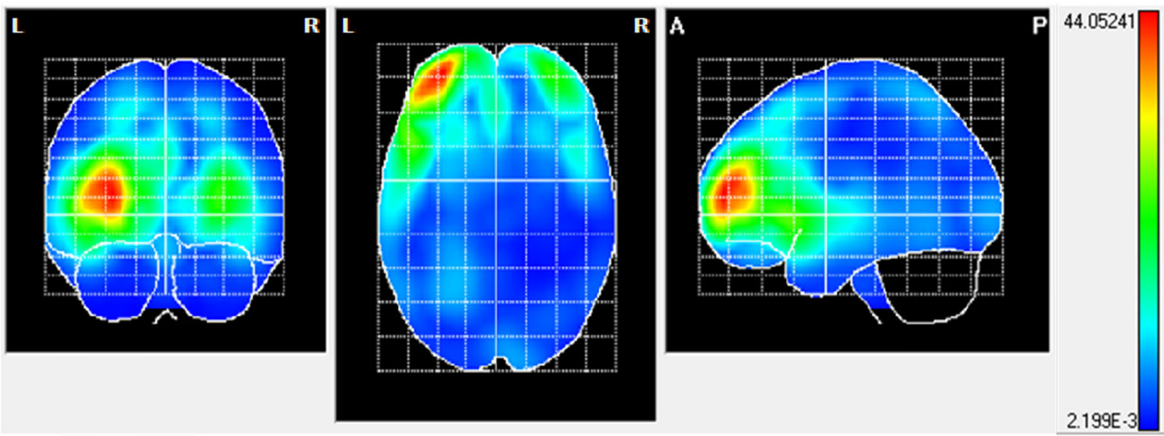

FIGURE 3. Mean electrical maps for Condition 2 (hearing impaired participants, muted video). Maximal intensity projection areas are displayed in yellow/red color. SPMs were computed based on a voxel-by-voxel Hotelling's T^2 test against zero.

TABLE 1. Condition 1.

AAL	X	Y	Z	Activation (T2)
Temporal_Sup_R	62	-18	-4	30.2320
Temporal_Mid_R	62	-18	-8	30.0220
Heschl_R	54	-10	4	20.8320
Temporal_Pole_Sup_R	54	2	0	20.7600
Temporal_Inf_R	62	-22	-20	20.6870
Insula_R	50	2	-4	20.2430
Rolandic_Oper_R	54	6	0	19.4980
Frontal_Inf_Oper_R	50	10	0	18.3190
Temporal_Pole_Mid_R	54	2	-16	17.8680
Frontal_Inf_Tri_R	50	22	0	14.5790

In the Condition 1 (hearing participants, video + music randomly assigned), as shown in Fig. 2 and Table 1, maximum statistically significant activations are obtained in the temporal lobe (superior, middle, and inferior areas, Heschl's area and temporal pole), followed by the insula, the Rolandic operculum, and the inferior frontal gyrus. All these activations are in the right hemisphere.

In the Condition 2 (hearing-impaired participants, muted video), as shown in Fig. 3 and Table 2, the maximum activation peaks are in the frontal areas. We also find less significant activation in the temporal pole. Most of the activations are found in the left hemisphere although right superior and medial frontal areas also show significant activations.

Finally, in the Condition 3 (hearing-impaired participants, muted video + vibrotactile stimuli), as shown in Fig. 4 and Table 3, extremely significant activations are obtained in the temporal lobe (superior and medial areas and temporal pole), the inferior frontal area, the insula, and the Rolandic operculum. Less significant peaks are also found in the medial frontal and Heschl's areas. All these activations are in the left hemisphere.

IV. DISCUSSION

The control group (condition 1, hearing participants, video + music randomly assigned) showed brain activations

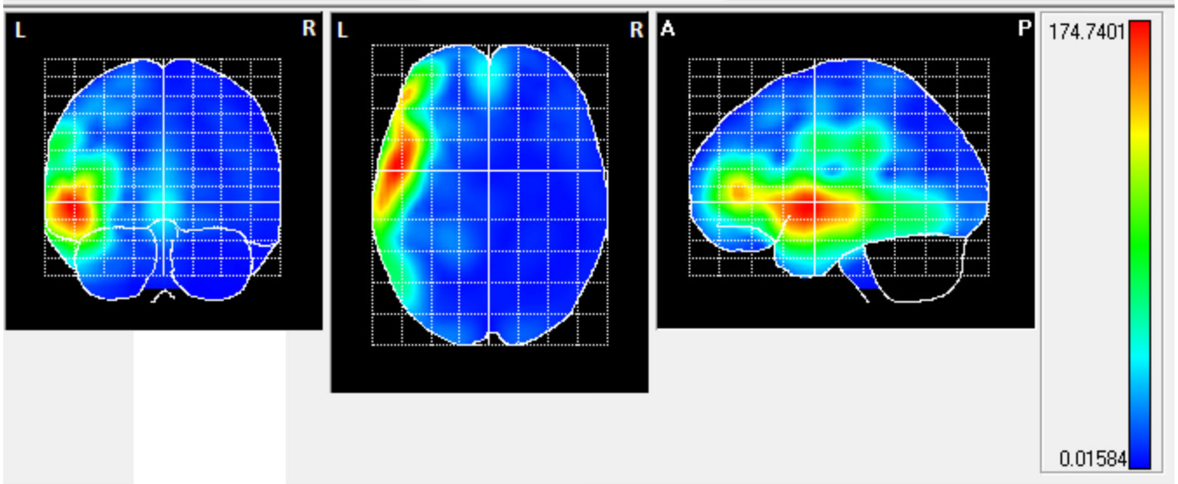

FIGURE 4. Mean electrical maps with Condition 3 (hearing-impaired participants, muted video + vibrotactile stimuli). Maximal intensity projection areas are displayed in yellow/red color. SPMs were computed based on a voxel-by-voxel Hotelling's T^2test against zero.

TABLE 2. Condition 2.

AAL	X	Y	Z	Activation (T2)
Frontal_Mid_L	-30	58	12	44.0520
Frontal_Sup_L	-26	58	12	39.5670
Frontal_Inf_Tri_L	-42	46	12	39.2800
Frontal_Sup_Orb_L	-30	58	-4	32.0470
Frontal_Mid_Orb_L	-34	58	-4	31.7150
Frontal_Inf_Orb_L	-46	18	-4	25.3910
Frontal_Sup_R	30	58	8	24.5120
Frontal_Mid_R	30	58	4	23.7370
Frontal_Inf_Oper_L	-50	14	0	23.2600
Temporal_Pole_Sup_L	-50	14	-4	22.5480

TABLE 3. Condition 3.

AAL	X	Y	Z	Activation (T2)
Temporal_Sup_L	-50	6	-4	174.7400
Temporal_Pole_Sup_L	-54	6	0	170.9190
Frontal_Inf_Oper_L	-50	10	0	167.2420
Rolandic_Oper_L	-54	6	4	161.4850
Insula_L	-46	10	-8	161.3440
Temporal_Mid_L	-58	-6	-8	157.1770
Frontal_Inf_Orb_L	-46	18	-4	156.2220
Frontal_Inf_Tri_L	-50	18	0	153.5690
Frontal_Mid_L	-46	46	0	132.2130
Heschl_L	-58	-10	8	104.0720

consistent with the existing literature, specifically the above-mentioned areas involved in affective music processing: superior temporal pole, Heschl's area (auditory cortex), insula and inferior frontal gyrus. In addition to these areas we found activation in medial temporal gyrus and Rolandic operculum.

The main peaks corresponded to the right superior temporal gyrus which has been identified as a region for multisensory integration [31], [32] and the medial temporal gyrus, which is associated with visual and auditory processing [4]. We found as well a greater activation of the Heschl's gyrus, which is part of the superior temporal lobe and contains the primary auditory cortex, and in the temporal pole, part of the association cortex, also involved in multimodal sensory integration [48], [49]. The insula, which also shows significant activation, is usually considered as a mediator between sensory and affective brain systems in the perception of affective sounds [50], translating the affective cue perception of sounds into subjective emotions [33], [50]. Functional imaging studies have shown activation of the insula during audio-visual integration tasks and passive listening to music [51], [52]. The activation of the Rolandic operculum has been related to the processing of pleasant music syntactic information [44]. Neutral images are used to reduce their emotional impact and, thus, make the measurement of the music and vibrotactile effects clearer, according to [41], [42].

In Condition 2 (hearing-impaired participants, muted videos), the activation was located mainly in the frontal lobe, associated with higher voluntary attentional resources and emotional and cognitive integration [53], [54]. The medial frontal cortex was significantly activated, which is involved in several social and emotional functions related to interpersonal communication and understanding [35]. This result was coherent with previous experiments with hearing-impaired participants who mobilize these areas with high intensity when watching a muted video [11].

In Condition 3 (hearing-impaired participants, muted video + vibrotactile stimuli), we again found peak activation in the same areas of condition 1, reported to be involved in affective music processing: superior temporal lobe, inferior frontal gyrus, Rolandic operculum, and insula, with similar significance level. The activations of the Heschl's area (auditory cortex), although with less significance, is an interesting result in hearing-impaired participants. Another significant area of activation was the medial frontal lobe, which is associated with visual and auditory processing [4]. The most remarkable difference with Condition 1 is the inversion of the laterality, probably because the hearing-impaired participants received the vibrotactile stimulation through their right hand and, hence, the left hemisphere would be the one in charge of processing it. The activation of the auditory cortex is in concordance with other studies showing that tactile vibrations activate the auditory cortex or improve musical and speech auditory perception [55]–[57].

The interesting finding in this study is that all the emotional neural networks involved in affective music processing are also significantly activated in the vibrotactile condition when combined with visual media. Although this study presents some limitations, as the size of the sample, single-sex participants, one-hand stimulation or narrow variability of ages of the control group, the obtained results encourage a deeper study of different vibrotactile stimulations that may show emotional arousal capabilities to enhance the visual experience of deaf and hard-of-hearing collectives.

Coin motors used in this study limit their operation signal to a certain frequency range. We have used a signal of 1 kHz triggered at the rhythm of the music. Other vibrational motors may work at different frequency ranges, producing different skin responses. Further research is needed to check the dependency of the results on the motors driving signals, or even using different ones for each finger.

V. CONCLUSION

We have found that vibrotactile stimuli can generate cortex activation in similar ways to sounds, in correspondence with the available literature. What has been shown specifically in our experiment is that the integration of very simple vibrotactile stimuli with neutral images can enhance the activation of auditory and emotional cortex areas, thus approaching the brain reactions to the ones of hearing subjects exposed to a complete audiovisual experience. This experiment has been set up as a proof of concept but, despite its limitations in terms of small participant number and the simplicity of music to vibrotactile mapping, these results open a wide path of research that seems to be able to provide new solutions for deaf and hard of hearing people, in the near future.

This work justifies further explorations of the touch channel, as an alternative to captions, to transmit the emotional information contained in the audio-visual soundtrack, and thus elicit the intended emotion in hearing-impaired subjects through vibrotactile stimulation.

ACKNOWLEDGMENT

The authors would like to thank the volunteers that participated in the experiment.

REFERENCES

[1] D. R. Brown and J. F. Cavanagh, "The sound and the fury: Late positive potential is sensitive to sound affect," *Psychophysiology*, vol. 54, no. 12, pp. 1812–1825, Dec. 2017.

[2] K. Donnelly, "The spectre of sound: Music in film and television," in *The Founders' Constitution*, vol. 1987, P. B. Kurland and R. Lerner, Eds. Chicago, IL, USA: Univ. Chicago Press, 2005. Accessed: Feb. 28, 2010. [Online]. Available: http://press-pubs.uchicago.edu/founders/

[3] W. F. Thompson, F. A. Russo, and D. Sinclair, "Effects of underscoring on the perception of closure in filmed events," *Psychomusicol., J. Res. Music Cognition*, vol. 13, nos. 1–2, pp. 9–27, 1994.

[4] C. Pehrs, L. Deserno, J.-H. Bakels, L. H. Schlochtermeier, H. Kappelhoff, A. M. Jacobs, T. H. Fritz, S. Koelsch, and L. Kuchinke, "How music alters a kiss: Superior temporal gyrus controls fusiform–amygdalar effective connectivity," *Social Cognit. Affect. Neurosci.*, vol. 9, no. 11, pp. 1770–1778, Nov. 2014.

[5] A. Gabrielsson, "Emotions in strong experiences with music," in *Music and Emotion: Theory and Research*. New York, NY, USA: Oxford Univ. Press, 2001, pp. 431–449.

[6] L.-L. Balkwill and W. F. Thompson, "A cross-cultural investigation of the perception of emotion in music: Psychophysical and cultural cues," *Music Perception*, vol. 17, no. 1, pp. 43–64, Oct. 1999.

[7] I. Molnar-Szakacs and K. Overy, "Music and mirror neurons: From motion to 'e'motion," *Social Cogn. Affect. Neurosci.*, vol. 1, no. 3, pp. 235–241, Dec. 2006.

[8] UN. (2006). *Convention on the Rights of Persons With Disabilities (CRPD)*. [Online]. Available: https://www.un.org/development/desa/disabilities/convention-on-the-rights-of-persons-with-disabilities.html

[9] BOE. (2010). *Ley 7/2010, de 31 de marzo, General de la Comunicación Audiovisual*. Accessed: Feb. 4, 2020. [Online]. Available: https://www.boe.es/buscar/pdf/2010/BOE-A-2010-5292-consolidado.pdf

[10] AENOR. (2012). *UNE 153010:2012. Subtitulado para personas sordas y personas con discapacidad auditiva*. Accessed: Feb. 4, 2020. [Online]. Available: https://www.une.org/encuentratu-norma/busca-tu-norma/norma/?c=N0049426

[11] P. Revuelta, T. Ortiz, M. J. Lucía, B. Ruiz, and J. M. Sánchez-Pena, "Limitations of standard accessible captioning of sounds and music for deaf and hard of hearing people: An EEG study," *Frontiers Integrative Neurosci.*, vol. 14, p. 1, Feb. 2020.

[12] L. A. Jones and S. J. Lederman, *Human Hand Function*. London, U.K.: Oxford Univ. Press, 2006.

[13] V. Hayward, "A brief overview of the human somatosensory system," in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Cham, Switzerland: Springer, 2018, ch. 3, pp. 29–72.

[14] C. Hopkins, S. Maté-Cid, R. Fulford, G. Seiffert, and J. Ginsborg, "Vibrotactile presentation of musical notes to the glabrous skin for adults with normal hearing or a hearing impairment: Thresholds, dynamic range and high-frequency perception," *PLoS ONE*, vol. 11, no. 5, May 2016, Art. no. e0155807.

[15] F. A. Russo, P. Ammirante, and D. I. Fels, "Vibrotactile discrimination of musical timbre.," *J. Exp. Psychol., Hum. Perception Perform.*, vol. 38, no. 4, pp. 822–826, 2012.

[16] G. W. Young, D. Murphy, and J. Weeter, "Vibrotactile discrimination of pure and complex waveforms," in *Proc. 12th Sound Music Comput. Conf.*, 2015, pp. 359–362.

[17] C. Kayser, C. I. Petkov, M. Augath, and N. K. Logothetis, "Studies from max Planck institute for biological cybernetics further understanding of magnetic resonance," *Health Med. Week*, vol. 48, no. 2, p. 2733, Sep. 2007.

[18] G. Caetano and V. Jousmäki, "Evidence of vibrotactile input to human auditory cortex," *NeuroImage*, vol. 29, no. 1, pp. 15–28, Jan. 2006.

[19] J. Huang, D. Gamble, K. Sarnlertsophon, X. Wang, and S. Hsiao, "Feeling music: Integration of auditory and tactile inputs in musical meter perception," *PLoS ONE*, vol. 7, no. 10, Oct. 2012, Art. no. e48496.

[20] S. Convento, K. A. Wegner-Clemens, and J. M. Yau, "Reciprocal interactions between audition and touch in flutter frequency perception," *Multisensory Res.*, vol. 32, no. 1, pp. 67–85, 2019.

[21] L. E. Crommett, A. Pérez-Bellido, and J. M. Yau, ''Auditory adaptation improves tactile frequency perception,'' *J. Neurophysiol.*, vol. 117, no. 3, pp. 1352–1362, Mar. 2017.

[22] M. S. Rahman, K. A. Barnes, L. E. Crommett, M. Tommerdahl, and J. M. Yau, ''Auditory and tactile frequency representations are co-embedded in modality-defined cortical sensory systems,'' *NeuroImage*, vol. 215, Jul. 2020, Art. no. 116837.

[23] T. Eerola, A. Friberg, and R. Bresin, ''Emotional expression in music: Contribution, linearity, and additivity of primary musical cues,'' *Frontiers Psychol.*, vol. 4, pp. 11–21, Jul. 2013.

[24] A. Gabrielsson and E. Lindström, ''The role of structure in the musical expression of emotions,'' in *Music and Emotion: Theory, Research, Applications*. Oxford, U.K.: Oxford Univ. Press, 2010, pp. 367–400.

[25] C. Branje and D. I. Fels, ''Playing vibrotactile music: A comparison between the vibrochord and a piano keyboard,'' *Int. J. Hum.-Comput. Stud.*, vol. 72, no. 4, pp. 431–439, Apr. 2014.

[26] R. Jack, A. McPherson, and T. Stockman, ''Designing tactile musical devices with and for deaf users: A case study,'' in *Proc. Int. Conf. Multimodal Exper. Music*, 2015, pp. 1–7.

[27] S. Holland, A. Bouwer, and O. Hödl, ''Haptics for the development of fundamental rhythm skills, including multi-limb coordination,'' in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Cham, Switzerland: Springer, 2018, ch. 11, pp. 215–237.

[28] S. Merchel and M. Ercan, ''Auditory-tactile experience of music,'' in *Musical Haptics*, S. Papetti and C. Saitis, Eds. Cham, Switzerland: Springer, 2018, ch. 7, pp. 123–148.

[29] E. Gunther and S. O'Modhrain, ''Cutaneous grooves: Composing for the sense of touch,'' *J. New Music Res.*, vol. 32, no. 4, pp. 369–381, Dec. 2003.

[30] S. Frühholz, W. Trost, and S. A. Kotz, ''The sound of emotions—Towards a unifying neural network perspective of affective sound processing,'' *Neurosci. Biobehav. Rev.*, vol. 68, pp. 96–110, Sep. 2016.

[31] B. Kreifelts, T. Ethofer, W. Grodd, M. Erb, and D. Wildgruber, ''Audiovisual integration of emotional signals in voice and face: An event-related fMRI study,'' *NeuroImage*, vol. 37, no. 4, pp. 1445–1456, Oct. 2007.

[32] V. Menon and D. J. Levitin, ''The rewards of music listening: Response and physiological connectivity of the mesolimbic system,'' *NeuroImage*, vol. 28, no. 1, pp. 175–184, Oct. 2005.

[33] S. A. Kotz, C. Kalberlah, J. Bahlmann, A. D. Friederici, and J.-D. Haynes, ''Predicting vocal emotion expressions from the human brain,'' *Hum. Brain Mapping*, vol. 34, no. 8, pp. 1971–1981, Aug. 2013.

[34] D.-E. Bamiou, F. E. Musiek, and L. M. Luxon, ''The insula (Island of Reil) and its role in auditory processing,'' *Brain Res. Rev.*, vol. 42, no. 2, pp. 143–154, May 2003.

[35] D. M. Amodio and C. D. Frith, ''Meeting of minds: The medial frontal cortex and social cognition,'' *Nature Rev. Neurosci.*, vol. 7, no. 4, pp. 268–277, Apr. 2006.

[36] A. J. Blood and R. J. Zatorre, ''Intensely pleasurable responses to music correlate with activity in brain regions implicated in reward and emotion,'' *Proc. Nat. Acad. Sci. USA*, vol. 98, no. 20, pp. 11818–11823, Sep. 2001.

[37] J. Panksepp and G. Bernatzky, ''Emotional sounds and the brain: The neuro-affective foundations of musical appreciation,'' *Behav. Processes*, vol. 60, no. 2, pp. 133–155, Nov. 2002.

[38] C. Lithari, C. A. Frantzidis, C. Papadelis, A. B. Vivas, M. A. Klados, C. Kourtidou-Papadeli, C. Pappas, A. A. Ioannides, and P. D. Bamidis, ''Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions,'' *Brain Topogr.*, vol. 23, no. 1, pp. 27–40, Mar. 2010.

[39] C. J. Cela-Conde, F. J. Ayala, E. Munar, F. Maestu, M. Nadal, M. A. Capo, D. del Rio, J. J. Lopez-Ibor, T. Ortiz, C. Mirasso, and G. Marty, ''Sex-related similarities and differences in the neural correlates of beauty,'' *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 10, pp. 3847–3852, Mar. 2009.

[40] BIAP. (1994). *BIAP Recommendation 02/1: Audiometric Classification of Hearing Impairments*. Accessed: Feb. 4, 2020. [Online]. Available: https://www.biap.org/en/recommandations/recommendations/tc-02-classification/213-rec-02-1-enaudiometric-classification-of-hearing-impairments/file

[41] A. B. M. Gerdes, M. J. Wieser, F. Bublatzky, A. Kusay, M. M. Plichta, and G. W. Alpers, ''Emotional sounds modulate early neural processing of emotional pictures,'' *Frontiers Psychol.*, vol. 4, p. 741, Oct. 2013.

[42] Y. Wang, M. Jiang, Y. Huang, Z. Sheng, X. Huang, W. Lin, Q. Chen, X. Li, Z. Luo, and B. Lv, ''Physiological and psychological effects of watching videos of different durations showing urban bamboo forests with varied structures,'' *Int. J. Environ. Res. Public Health*, vol. 17, no. 10, p. 3434, May 2020.

[43] C. L. Krumhansl, ''An exploratory study of musical emotions and psychophysiology,'' *Can. J. Exp. Psychol./Revue canadienne de psychologie expérimentale*, vol. 51, no. 4, pp. 336–353, 1997.

[44] S. Koelsch, S. Skouras, T. Fritz, P. Herrera, C. Bonhage, M. B. Küssner, and A. M. Jacobs, ''The roles of superficial amygdala and auditory cortex in music-evoked fear and joy,'' *NeuroImage*, vol. 81, pp. 49–60, Nov. 2013.

[45] H. Oey and V. Mellert, ''Vibration thresholds and equal vibration levels at the human fingertip and palm,'' in *Proc. 5th Int. Congr. Acoust.*, 2004, pp. 1–4.

[46] R. D. Pascual-Marqui, C. M. Michel, and D. Lehmann, ''Low resolution electromagnetic tomography: A new method for localizing electrical activity in the brain,'' *Int. J. Psychophysiol.*, vol. 18, no. 1, pp. 49–65, Oct. 1994.

[47] A. C. Evans, D. L. Collins, S. R. Mills, E. D. Brown, R. L. Kelly, and T. M. Peters, ''3D statistical neuroanatomical models from 305 MRI volumes,'' in *Proc. IEEE Conf. Rec. Nucl. Sci. Symp. Med. Imag. Conf.*, Oct. 1993, pp. 1813–1817.

[48] I. R. Olson, A. Plotzker, and Y. Ezzyat, ''The enigmatic temporal pole: A review of findings on social and emotional processing,'' *Brain*, vol. 130, no. 7, pp. 1718–1731, May 2007.

[49] L. M. Skipper, L. A. Ross, and I. R. Olson, ''Sensory and semantic category subdivisions within the anterior temporal lobes,'' *Neuropsychologia*, vol. 49, no. 12, pp. 3419–3429, Oct. 2011.

[50] F. Mirz, A. Gjedde, H. Sdkilde-Jrgensen, and C. B. Pedersen, ''Functional brain imaging of tinnitus-like perception induced by aversive auditory stimuli,'' *NeuroReport*, vol. 11, no. 3, pp. 633–637, Feb. 2000.

[51] D. Wildgruber, ''Distinct frontal regions subserve evaluation of linguistic and emotional aspects of speech intonation,'' *Cerebral Cortex*, vol. 14, no. 12, pp. 1384–1389, May 2004.

[52] P. J. Lang and M. M. Bradley, ''Emotion and the motivational brain,'' *Biol. Psychol.*, vol. 84, no. 3, pp. 437–450, Jul. 2010.

[53] K. O. Bushara, T. Hanakawa, I. Immisch, K. Toma, K. Kansaku, and M. Hallett, ''Neural correlates of cross-modal binding,'' *Nature Neurosci.*, vol. 6, no. 2, pp. 190–195, Feb. 2003.

[54] K. B. E. Böcker, J. M. P. Baas, J. L. Kenemans, and M. N. Verbaten, ''Stimulus-preceding negativity induced by fear: A manifestation of affective anticipation,'' *Int. J. Psychophysiol.*, vol. 43, no. 1, pp. 77–90, Dec. 2001.

[55] M. Schürmann, G. Caetano, Y. Hlushchuk, V. Jousmäki, and R. Hari, ''Touch activates human auditory cortex,'' *NeuroImage*, vol. 30, no. 4, pp. 1325–1331, May 2006.

[56] S. B. Leder, J. B. Spitzer, P. Milner, C. Flevaris-Phillips, and F. Richardson, ''Vibrotactile stimulation for the adventitiously deaf: An alternative to cochlear implantation,'' *Arch. Phys. Med. Rehabil.*, vol. 67, no. 10, pp. 754–758, Oct. 1986.

[57] X. Luo and L. Hayes, ''Vibrotactile stimulation based on the fundamental frequency can improve melodic contour identification of normal-hearing listeners with a 4-channel cochlear implant simulation,'' *Frontiers Neurosci.*, vol. 13, p. 1145, Oct. 2019.

MARÍA J. LUCÍA was born in Ponferrada, León, Spain, in 1962. She received the M.S. degree in psychology and the M.S. degree in physics from the Complutense University of Madrid, in 1985 and 1989, respectively. She is currently pursuing the Ph.D. degree in computer sciences with the Carlos III University of Madrid.

From 1989 to 2016, she worked with Telefónica España as the Consultant and Manager for ICT projects. In 2017, she joined the Carlos III University of Madrid, where she is currently an Associate Professor and collaborates with the Spanish Center for Captioning and Audio Description (CESyA) as a Researcher.

PABLO REVUELTA was born in Madrid, Spain, in 1980. He received the B.S. degree in telecommunications engineering, the M.S. degree in advanced electronics, and the Ph.D. degree in electronic engineering from the Carlos III University of Madrid, in 2006, 2008, and 2013, respectively.

Since 2006, he has been a Researcher with the Displays and Photonic Applications Group, Carlos III University, the Spanish Center for Captioning and Audio Description (CESyA), and the start-up Visión Táctil S. L. He also worked with the Digital Signal Processing Lab 5 (EPFL, Switzerland) in 2010 and the Georgia Tech Sonification Laboratory, USA, in 2012. He has been teaching with the Oviedo University since 2018. He is currently with CESyA in Madrid, Spain. He has authored three international book chapters, 20 conference communications, six journal reports, and one patent.

ÁLVARO GARCÍA was born in Madrid, Spain, in 1986. He received the B.S. degree in industrial electronics technical engineering and the M.S. degree in industrial electronics from the Carlos III University of Madrid, in 2010 and 2016, respectively, where he is currently pursuing the Ph.D. degree in electric, electronic, and automatic engineering.

He currently combines work in a private electronics company with research with the Carlos III University of Madrid.

BELÉN RUIZ was born in Caracas, Venezuela, in 1959. She received the M.S. degree in physics from the Complutense University of Madrid, in 1983, and the Ph.D. degree in computer science from the Carlos III University of Madrid, in 1998.

From 1984 to 1996, she worked with ICT companies, such as ALCATEL and INDRA. In 1996, she joined the Carlos III University of Madrid, where she has been teaching as a Professor with the Department of Informatics, since 2018, and the Dean of the Computer Science Degree. She is currently the Director of the Spanish Center for Captioning and Audio Description (CESyA) and a Co-Director of both the disability-EADS-Adecco Foundation UC3M and the MA TADIS. She has led several national and international research projects on voice recognition, human–machine interface, software engineering, systems analysis, mobile communications, and audiovisual accessibility. She has written more than 30 reports in the field of biometrics and the application of technology to help people with disabilities. She is the author of nine patents, six SW records, and six trademark registrations.

RICARDO VERGAZ received the M.S. degree in physics and electronics engineering and the Ph.D. degree in science (physics) from the University of Valladolid, Valladolid, Spain, in 2001.

Since 2001, he has been with the Universidad Carlos III de Madrid, Leganés, Spain, where he is currently an Associate Professor with the Displays and Photonics Applications Group, Department of Tecnología Electrńica. He has coauthored more than 50 JCR-indexed publications and 70 communications in international conferences. His current research interests include electro-optical materials and devices characterization, design of novel nanometric devices using scattering directionality, and technical aids for different impairments.

VÍCTOR CERDÁN was born in Madrid, Spain, in 1983. He received the M.S. degree in journalism and the Ph.D. degree in information sciences from the Complutense University of Madrid in 2006 and 2011, respectively.

He is currently an Assistant Professor with the Applied Communication Sciences Department, Complutense University of Madrid. In the professional field, he has directed two seasons of the series of international documentaries Héroes Invisibles for Spanish public television (RTVE). In cinema, he has directed and produced the short films: Radio Atacama, El Mal, and Caracoles Serranos, awarded at international film festivals and exhibited in: Hamburg, Germany, Vila Do Conde, Portugal, Montreal, Canada, Florida, USA, Munich, Germany, Mexico, Mexico, Barcelona, Spain, Paris, France, Shenzhen, China, Seoul, South Korea, and Brest, France. He has written four book chapters and ten reports in the fields of communication and social sciences.

TOMÁS ORTIZ was born in Sacecorbo, Guadalajara, Spain, in 1950. He received the M.S. degree in psychology, the M.S. degree in medicine, the Ph.D. degree in psychology, and the Ph.D. degree in Medicine from the Complutense University of Madrid in 1975, 1988, 1980, and 1996, respectively.

He has been teaching neuropsychology, medical psychology, and psychiatry, for more than 30 years, and has tutored more than 30 Ph.D. theses in these areas. He is currently working with the Complutense University of Madrid, where he has been a Professor with the Psychiatry Department, since 1999. He also works with the start-up Vision Táctil S. L., as well as in other projects related to neuroscience and education. He has published 220 reports in both national and international journals. He is a member of the Royal National Academy of Medicine since 2016.